

パターン認識 レポート 2

1019163 日置竜輔 *

2021 年 7 月 26 日

今回は weka を使ってスポーツをする人に身長と運動の好き嫌いが関係してるかを調べるために、以下のモデルを作成した。

Listing 1 今回 weka を用いて 3*3 で学習させるモデル

```
1 @relation play-sports
2
3 @attribute height {high, medium, low}
4 @attribute like {like, soso, dislike}
5 @attribute play-sports {-1, 1}
6
7 @data
8 high, like, 1
9 high, soso, 1
10 high, dislike, -1
11 medium, like, 1
12 medium, soso, -1
13 medium, dislike, -1
14 low, like, 1
15 low, soso, -1
16 low, dislike, -1
```

@attribute height は身長が「高い・平均・低い」の 3 段階で判断し、

@attribute like は運動が「好き・どちらでもない・嫌い」の 3 段階で判断し

@attribute play-sports で -1 ならば運動が嫌い、1 ならば運動が好きというモデルを作成した。

* 公立はこだて未来大学 システム情報科学部 複雑系知能学科 複雑系コース B3

@data は例えば、1 行目は身長が高く、運動が好きならばスポーツをしている、9 行目は身長が低く、運動が嫌いならばスポーツをしていないといったモデルを作成した。

weka を用いて上記のデータを学習を行った結果、以下の結果が得られた。

Listing 2 weka を用いて上記のデータを学習させた結果

```
1 === Run information ===
2
3 Scheme: weka.classifiers.functions.SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 2 -V -1 -W 1 -K
   "weka.classifiers.functions.supportVector.NormalizedPolyKernel -E 3.0 -C 250007" -
   calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places
   4"
4 Relation: play-sports
5 Instances: 9
6 Attributes: 3
7           height
8           like
9           play-sports
10 Test mode: evaluate on training data
11
12 === Classifier model (full training set) ===
13
14 SMO
15
16 Kernel used:
17   Normalized Poly Kernel:  $K(x,y) = \langle x,y \rangle^3 / (\langle x,x \rangle^3 \langle y,y \rangle^3)^{1/2}$ 
18
19 Classifier for classes: -1, 1
20
21 BinarySMO
22
23   0.922 * <1 0 0 1 0 0 > * X]
24 + 1 * <0 0 1 1 0 0 > * X]
25 - 0.6378 * <0 1 0 0 0 1 > * X]
26 + 1 * <0 1 0 1 0 0 > * X]
27 - 0.6393 * <0 0 1 0 0 1 > * X]
28 - 0.8757 * <0 0 1 0 1 0 > * X]
29 - 0.8747 * <0 1 0 0 1 0 > * X]
30 - 0.8945 * <1 0 0 0 0 1 > * X]
31 + 1 * <1 0 0 0 1 0 > * X]
32 - 0.1854
33
```

```

34 Number of support vectors: 9
35
36 Number of kernel evaluations: 45 (96.815% cached)
37
38
39
40 Time taken to build model: 0.03 seconds
41
42 === Evaluation on training set ===
43
44 Time taken to test model on training data: 0 seconds
45
46 === Summary ===
47
48 Correctly Classified Instances 9 100 %
49 Incorrectly Classified Instances 0 0 %
50 Kappa statistic 1
51 Mean absolute error 0
52 Root mean squared error 0
53 Relative absolute error 0 %
54 Root relative squared error 0 %
55 Total Number of Instances 9
56
57 === Detailed Accuracy By Class ===
58
59          TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
60          1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 -1
61          1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 1
62 Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000
63
64 === Confusion Matrix ===
65
66  a b <-- classified as
67  5 0 | a = -1
68  0 4 | b = 1

```

14-17 行目では3次の多項式カーネルを利用していることが確認でき、23-32 行目で実行結果が正確に得られていることが確認できた。

weka を使用すると、モデルを作成することによってしっかりとした結果が具体的に得られることがわかった。

しかし、今回は作成したモデルが少なかったため、さらにデータを学習させれば精度の高い結果が得られるのではないかと考えた。